

## Observed Relationships between Arctic Longwave Cloud Forcing and Cloud Parameters Using a Neural Network

YONGHUA CHEN\*

*Institute of Marine and Coastal Sciences, Rutgers—The State University of New Jersey, New Brunswick, New Jersey*

FILIPPE AIRES

*Laboratoire de Météorologie Dynamique du CNRS/Institut Pierre Simon Laplace, Université Pierre et Marie Curie, Paris, France*

JENNIFER A. FRANCIS AND JAMES R. MILLER

*Institute of Marine and Coastal Sciences, Rutgers—The State University of New Jersey, New Brunswick, New Jersey*

(Manuscript received 20 June 2005, in final form 27 October 2005)

### ABSTRACT

A neural network technique is used to quantify relationships involved in cloud–radiation feedbacks based on observations from the Surface Heat Budget of the Arctic (SHEBA) project. Sensitivities of longwave cloud forcing (CFL) to cloud parameters indicate that a bimodal distribution pattern dominates the histogram of each sensitivity. Although the mean states of the relationships agree well with those derived in a previous study, they do not often exist in reality. The sensitivity of CFL to cloud cover increases as the cloudiness increases with a range of  $0.1\text{--}0.9\text{ W m}^{-2}\text{ \%}^{-1}$ . There is a saturation effect of liquid water path (LWP) on CFL. The highest sensitivity of CFL to LWP corresponds to clouds with low LWP, and sensitivity decreases as LWP increases. The sensitivity of CFL to cloud-base height (CBH) depends on whether the clouds are below or above an inversion layer. The relationship is negative for clouds higher than 0.8 km at the SHEBA site. The strongest positive relationship corresponds to clouds with low CBH. The dominant mode of the sensitivity of CFL to cloud-base temperature (CBT) is near zero and corresponds to warm clouds with base temperatures higher than  $-9^{\circ}\text{C}$ . The low and high sensitivity regimes correspond to the summer and winter seasons, respectively, especially for LWP and CBT. Overall, the neural network technique is able to separate two distinct regimes of clouds that correspond to different sensitivities; that is, it captures the nonlinear behavior in the relationships. This study demonstrates a new method for evaluating nonlinear relationships between climate variables. It could also be used as an effective tool for evaluating feedback processes in climate models.

### 1. Introduction

Clouds play an important role in the Arctic climate, especially in the ice–albedo and cloud–radiation feedback mechanisms. These feedbacks are believed to be responsible for polar amplification of global warming, and make the Arctic the most sensitive area to global climate change. In the past several decades, the Arctic has been undergoing dramatic changes in every aspect

of the system. Many of the changes observed during recent decades in the oceanic and terrestrial northern high latitudes are summarized in Serreze et al. (2000). Significant warming has occurred in the central Arctic, as have downward trends in sea ice cover and negative snow anomalies over both the North American and Eurasian continents. Arctic cloudiness has also changed substantially in the last two decades according to analyses of satellite-derived datasets (Wang and Key 2003; Comiso 2003; Schweiger 2004). While these analyses are not in complete agreement, it appears that cloud amount has decreased during winter and increased during spring and autumn.

Our understanding of cloud–radiation interactions and feedbacks in the Arctic is, however, still limited by data sparsity and/or by poor spatial and temporal sam-

---

\* Current affiliation: Department of Applied Physics and Mathematics, Columbia University, New York, New York.

---

Corresponding author address: Yonghua Chen, NASA GISS, 2880 Broadway, New York, NY 10025.  
E-mail: ychen@giss.nasa.gov

pling. Recently available Arctic data from the Surface Heat Budget of the Arctic (SHEBA) ocean project offer new opportunities to evaluate relationships between cloud properties and radiative forcing (e.g., Shupe and Intrieri 2004). Even though the record length and spatial representation are limited, these observations are believed to be some of the most accurate and comprehensive measurements with a high temporal resolution in the Arctic.

A quantitative evaluation of important feedback loops is difficult because of the many interactions among the relevant climate variables, particularly the nonlinear behavior of these interactions. In this paper, a neural network (NN) approach is pursued to capture some of these nonlinear relationships. The neural network method is an alternative statistical method to traditional regression techniques; it has been widely used in environmental science and water resources since the 1990s. For example, ozone concentration forecasts using a neural network have been studied by Yi and Prybutok (1996) and Gardner and Dorling (2001). Another area in which a neural network has been used extensively is in the retrievals of geophysical parameters from remotely sensed data (e.g., Davis et al. 1993; Thiria et al. 1993; Escobar et al. 1993). Neural network approximations are also used in numerical models to replace some parameterizations of physical processes to improve computational efficiency (e.g., Krasnopolsky and Chevallier 2003; Key and Schweiger 1998). In most applications, a feed-forward neural network is used; that is, information is processed only in one direction from input to hidden to output layers. Hornik et al. (1989) has shown that any smooth measurable function can be approximated by a neural network with one or more hidden layers. Most previous applications are focused on the direct use of output from the neural network, such as converting remotely sensed brightness temperatures to physical parameters.

A by-product of a neural network model, the NN Jacobian matrix, can be obtained from a trained neural network. This Jacobian matrix is first examined in an effort to understand the nature of the neural network beyond the standard “black box” level. These Jacobians have been used to add constraints in a radiative transfer model (e.g., Aires et al. 1999). The Jacobians can also be used for variational assimilation applications (Chevallier and Mahfouf 2001). The NN Jacobian has been used to investigate sensitivities in a remote sensing algorithm (Aires et al. 2001). In a totally different context, the NN Jacobians were put into a theoretical framework as a tool to study the instantaneous, multivariate, nonlinear sensitivities in climate feedback processes (Aires and Rossow 2003). Aires et al. (2004)

further investigated the uncertainties in the Jacobian matrix of multivariate sensitivities and proposed a principal component analysis regularization scheme. A recent review paper on cloud feedbacks by Stephens (2005) summarizes some of the major obstacles to our understanding of cloud feedbacks. He recommends extending the classical feedback diagnostics to investigate instantaneous sensitivities instead of equilibrium estimates. These sensitivities constitute a step toward a more realistic representation and evaluation of feedback processes, particularly in their time evolution and their roles in governing cloud–radiation interactions. In his review, he recommends the Aires and Rossow (2003) method as one way to obtain the instantaneous sensitivities, that is, to apply a neural network approach and examine the NN Jacobians.

The principal objective of this paper is to demonstrate the capability of a neural network to quantify nonlinear relationships between climate variables. To keep the analysis simple while demonstrating this technique, we will focus on bivariate cases, that is, examining the sensitivity between longwave cloud forcing (CFL) and each of four cloud parameters through an annual cycle at an Arctic location. The Arctic, with its dramatic annual swing between two very different regimes, is the type of region where traditional linear regression may founder if applied to a full year of observations. If the NN can capture these relationships, it will offer a new tool to investigate feedback processes because such sensitivities are the controlling factors for these processes. Quantified relationships will help us to improve our understanding of those feedbacks and future climate change.

We use the SHEBA dataset to examine relationships between cloud parameters and longwave cloud radiative forcing obtained from the neural network. We present these bivariate relationships as the Jacobians from the neural network, which are the first partial derivatives of CFL to each cloud parameter. Our NN-derived sensitivities between CFL and other cloud parameters are compared with those from Shupe and Intrieri (2004), who analyzed measurements of these variables from the SHEBA experiment. A description of the data is given in section 2. The neural network approach is introduced in section 3 with a simple test case. The results from the bivariate neural network analysis are given in section 4, followed by a summary and discussion in section 5.

## 2. Data description

The SHEBA data used in this paper were obtained from the Environmental Technology Laboratory (ETL) at hourly temporal resolution. They cover a pe-

riod from 1 November 1997 to 1 October 1998, and thus autumn conditions are underrepresented in the dataset. The variables considered here are CFL, cloud cover (CLD), column liquid water path (LWP), cloud-base height (CBH), and cloud-base temperature (CBT). The CFL is calculated as the difference between measured all-sky net longwave fluxes and modeled clear-sky longwave fluxes at the surface (Intrieri et al. 2002b). The other four variables are determined from direct measurements of radar or lidar during the SHEBA field experiment. In this section, a brief summary of each variable is given. Details about this dataset are presented in Intrieri et al. (2002a,b).

The CFL ranges from  $-20$  to  $80 \text{ W m}^{-2}$  with a bimodal distribution; one mode is centered near  $5 \text{ W m}^{-2}$  (mainly in winter), and the other near  $65 \text{ W m}^{-2}$  (Fig. 1a). The negative CFL is introduced by errors in modeling clear-sky fluxes, measurement errors, and instrument mismatches (Shupe and Intrieri 2004). Negative CFL values mainly correspond to conditions of clear sky, low LWP, low CBH, and cold CBT in winter. In winter and spring, small CFL values are common, while large values are more common during summer and the transition seasons.

The definition of cloud cover in this dataset is different from the traditional one. It is referred to as vertical cloud fraction, which is the percentage of time that the cloud sensor (lidar or radar) detects the occurrence of cloud directly overhead. The cloud cover used in this study is averaged over 1 h. The distribution of CLD during SHEBA is strongly bimodal (Fig. 1b): nearly 70% of the observations have cloud cover greater than 95% (overcast), and 20% have cloud cover less than 5% (clear). In winter, there are more clear cases than overcast ones, and vice versa in warm seasons. Cases with few clouds are characterized by small CFL, very low LWP, low CBH, and cold CBT. Mostly cloudy cases are responsible for large CFL during warm seasons, and warm CBT ( $>-10^\circ\text{C}$ ).

The LWP has a strong exponential distribution, with more than 80% of the LWP measurements being less than  $100 \text{ g m}^{-2}$  (Fig. 1c). High LWPs occur mainly in summer. There are a few spikes in the LWP retrievals; approximately 2% of the data are  $>300 \text{ g m}^{-2}$  and occur mostly in summer with high CBT and large CFL. We suspect that these extremely high LWPs are due to precipitation. There are also quite a few negative LWP values mainly in winter when the Arctic air is cold and dry. These are associated with errors in the retrieval algorithm. According to Westwater et al. (2001), the retrieval uncertainty is about  $25 \text{ g m}^{-2}$ ; thus, those below  $-25 \text{ g m}^{-2}$  can be treated as missing values, and are eliminated in the following analysis.

The CBH also has a strong exponential distribution, with more than 80% of the cloud bases below 1 km (Fig. 1d). Although there is no distinct seasonal dependence in CBH (Shupe and Intrieri 2004), there is some indication that clouds may be higher in winter and spring, except January, and lower in summer. Base heights near zero occur mainly under clear-sky conditions, and correspond to small CFL, low LWP, and cold CBT during winter and the transition seasons. Diamond dust, that is, precipitating small, unbranched ice crystals, is included in the cloud-base measurements.

The CBT has a significant impact on downward longwave flux, and consequently the CFL. Its distribution has a long tail toward the cold end with a sharp peak around  $-2^\circ\text{C}$ , and a broad one centered near  $-25^\circ\text{C}$  (Fig. 1e). Most values are between  $-35^\circ$  and  $1^\circ\text{C}$ . The cloud bases have higher temperatures in summer and autumn compared with those in winter and spring. Clouds with base temperatures higher than  $-15^\circ\text{C}$  mainly occur in summer and spring under overcast conditions and correspond to large CFL and low CBH. Extremely cold CBTs ( $<-30^\circ\text{C}$ ) occur mainly in winter and are associated with small CFL and LWP. These conditions occur with either very low cloud or very high cloud.

### 3. Neural network approach

#### a. Neural network

An NN is a powerful statistical model that is widely used in classification, pattern recognition, regression, and other scientific areas. In contrast to traditional regression, a fitting function is not assumed in an NN approach. Most NN applications focus on the direct output of the NN. In this paper, we are interested not only in these outputs, but also in the Jacobian matrix within the neural network, which contains the first partial derivatives of a given output variable with respect to a given input variable. This, by definition, is the sensitivity of CFL (output variable) to cloud parameters (input variables) inferred by the NN model. For example, if CFL is the output variable of a neural network, and LWP and CBT are the input variables, the Jacobians from this NN will contain  $\partial\text{CFL}/\partial\text{LWP}$  and  $\partial\text{CFL}/\partial\text{CBT}$ .

Figure 2 shows a typical neural network with a multilayered perceptron including one hidden layer. The three layers are connected by neuronal links with weights  $W$ . After assigning initial random values to the weights, the training process iterates to find a set of weights that minimize an error function  $E$ . The weights move in the direction of the negative gradient of the

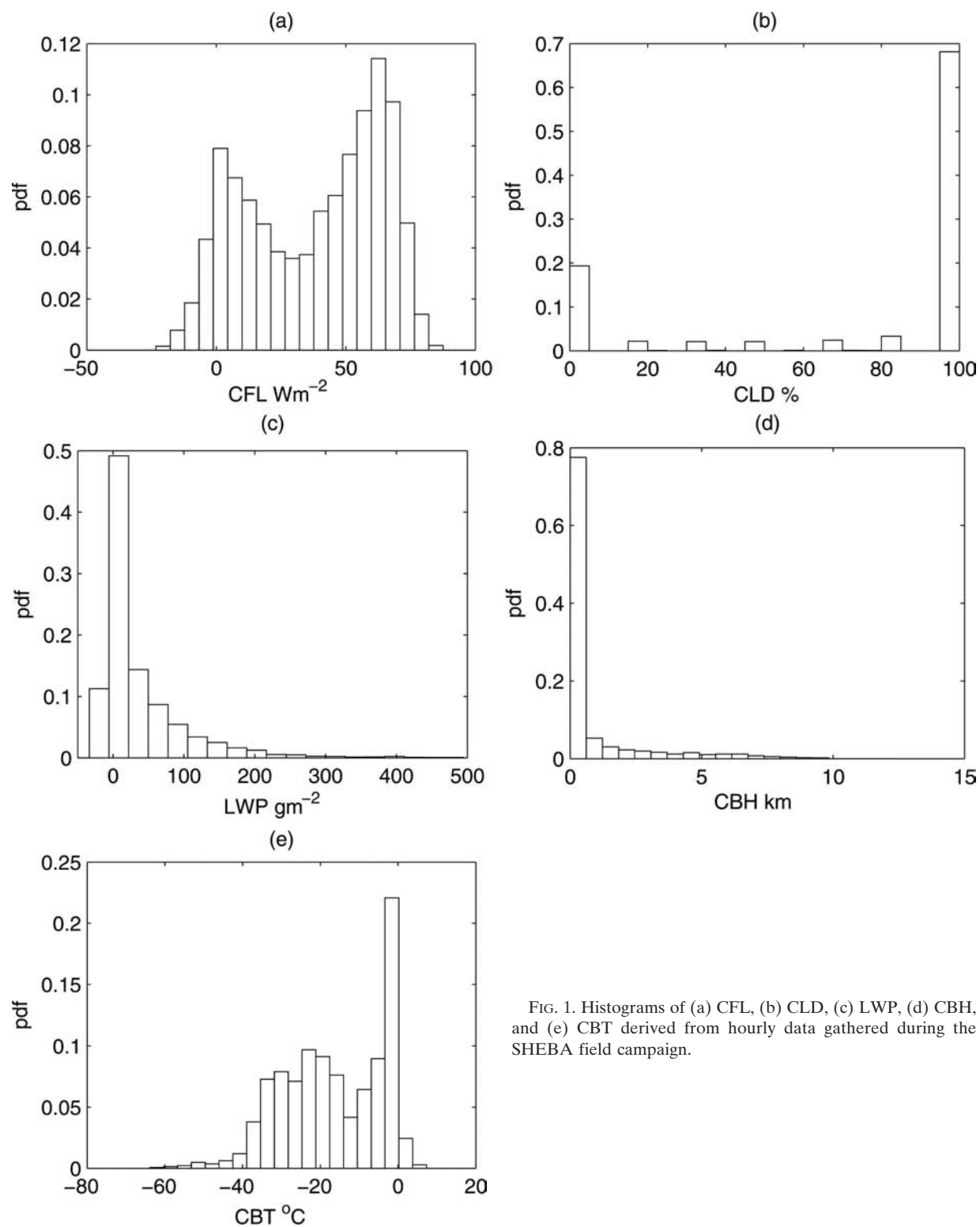


FIG. 1. Histograms of (a) CFL, (b) CLD, (c) LWP, (d) CBH, and (e) CBT derived from hourly data gathered during the SHEBA field campaign.

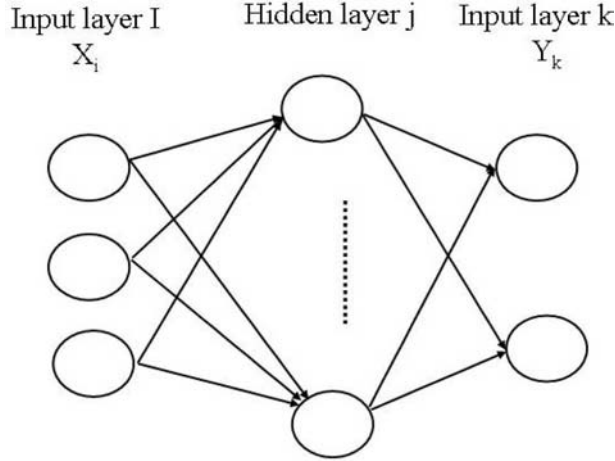


FIG. 2. Sketch of a feed-forward neural network with one hidden layer:  $i, j, k$  are indices for input-, hidden-, and output-layer neurons.

error function  $\Delta W = -\eta \nabla E$ , where  $\eta$  is the learning rate. During training, the weights of each link are estimated by an online (or stochastic) gradient descent algorithm; that is, the weights are updated immediately after incorporating each data point.

The activation function for the hidden layer is the hyperbolic tangent function,  $\sigma(a) = \tanh(a)$ . In matrix notation, the input matrix is  $\mathbf{X}_{n \times i}$  and the output matrix is  $\mathbf{Y}_{n \times k}$ , where  $n$  is sample size,  $i$  is number of input variables, and  $k$  is number of output variables. Suppose the matrix of weights that links input and hidden layers is  $\mathbf{W}_1$ , and the matrix that links hidden and output layers is  $\mathbf{W}_2$ , then the Jacobian matrix  $\mathbf{J}$  for the NN input  $x$  [ $x = (x_1, \dots, x_i)$  is one of the  $n$  samples in  $\mathbf{X}_{n \times i}$ ] is given by  $\mathbf{J}(x) = \mathbf{W}_2^T \sigma'(\mathbf{W}_1 \cdot \mathbf{x}) \mathbf{W}_1^T$ , where superscript T stands for transpose and the prime indicates derivative. For example, for  $i = 3$  and  $k = 2$ , we will have the  $2 \times 3$  Jacobian matrix  $\mathbf{J}(x)$ :

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_1}(x) & \frac{\partial y_1}{\partial x_2}(x) & \frac{\partial y_1}{\partial x_3}(x) \\ \frac{\partial y_2}{\partial x_1}(x) & \frac{\partial y_2}{\partial x_2}(x) & \frac{\partial y_2}{\partial x_3}(x) \end{bmatrix}.$$

The advantage of this neural network Jacobian is that it gives a direct statistical evaluation of the multivariate and nonlinear sensitivities that depends on each configuration of input and output variables (Aires and Rossow 2003). Aires et al. (2004) also proposed a regularization technique using a principal component analysis to suppress the multicollinearities to obtain robust Jacobians for multivariate cases. The results presented in this paper are for bivariate cases only.

#### b. Test case with specified functional relationships

Tests are performed in this section to examine the ability of a feed-forward neural network with one hidden layer to capture known relationships between variables, in particular the Jacobians of such relations. A simple example is given in this section for some nonlinear relationships. The neural network estimates are compared with those from the linear regression. The purpose of this section is not to validate theoretically this neural network approach, but rather to demonstrate that the NN can capture the nonlinear relationships much better than the traditional linear regression method. A more sophisticated theoretical case study can be found in Aires and Rossow (2003) using the Lorenz model.

The inputs of the multilayered perceptron with one hidden layer are given by  $x = (x_1, x_2, x_3)$ , where the three coordinates are random variables with normal distributions with different means and variances, that is,  $x_1 \sim N(4, 1)$  (with a mean of 4 and variance equal to 1),  $x_2 \sim N(0, 1)$ ,  $x_3 \sim N(0, 4)$ . The three variables are independent with a sample size of 8000. The results are valid for a variety of distributions. Three other variables,  $y_1, y_2$ , and  $y_3$ , are constructed according to some known arbitrary functions; that is,

$$y_1 = x_1^2 + x_2,$$

$$y_2 = \exp(x_2),$$

and

$$y_3 = x_1 x_2 + x_3.$$

Thus, the Jacobian matrix is given as

$$\begin{aligned} \mathbf{J}(x) &= \begin{bmatrix} \mathbf{J}_{11}(x) & \mathbf{J}_{12}(x) & \mathbf{J}_{13}(x) \\ \mathbf{J}_{21}(x) & \mathbf{J}_{22}(x) & \mathbf{J}_{23}(x) \\ \mathbf{J}_{31}(x) & \mathbf{J}_{32}(x) & \mathbf{J}_{33}(x) \end{bmatrix}, \\ &= \begin{bmatrix} \frac{\partial y_1}{\partial x_1}(x) & \frac{\partial y_1}{\partial x_2}(x) & \frac{\partial y_1}{\partial x_3}(x) \\ \frac{\partial y_2}{\partial x_1}(x) & \frac{\partial y_2}{\partial x_2}(x) & \frac{\partial y_2}{\partial x_3}(x) \\ \frac{\partial y_3}{\partial x_1}(x) & \frac{\partial y_3}{\partial x_2}(x) & \frac{\partial y_3}{\partial x_3}(x) \end{bmatrix}, \\ &= \begin{bmatrix} 2x_1 & 1 & 0 \\ 0 & \exp(x_2) & 0 \\ x_2 & x_1 & 1 \end{bmatrix}. \end{aligned}$$

Data are randomly divided into three parts for the purposes of training, cross validation, and testing. A neural network is optimized using the training dataset. The root-mean-square (RMS) error is monitored using both

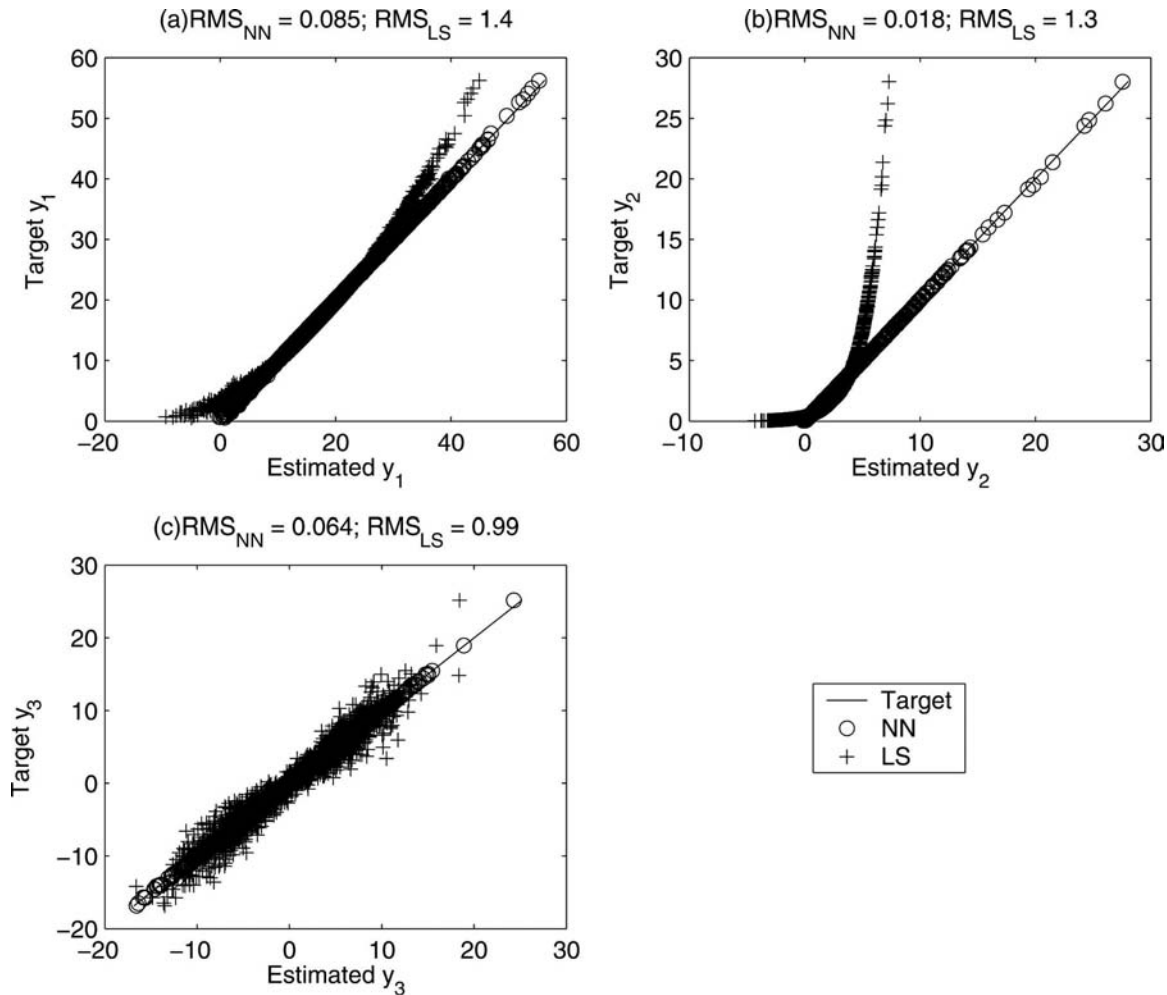


FIG. 3. Scatterplots comparing true values of each output variable with its estimates from the neural network and linear regression methods. The solid line represents the true values, circles represent the neural network estimate, and crosses represent the linear regression estimate.

training and cross-validation data, and the training process is stopped when the RMS error between iterations is small, or when the RMS error for the cross-validation data starts increasing, which is referred to as the early stopping technique (Bishop 1996). The test dataset is used to perform an independent check on the trained neural network.

The fitting results for each of the output variables are plotted in Fig. 3 for both linear regression (crosses) and neural network (circles). Given that the RMS errors for each estimate of  $y_1$ ,  $y_2$ , and  $y_3$  using both methods are much less than the standard deviation for  $y_1$ ,  $y_2$ , and  $y_3$  (8.01, 2.16, and 4.60), respectively, both methods can be used, but the neural network estimate performs better than does the linear regression. The fitting bias from the neural network is smaller than that from the linear regression, and more variability is explained by the

neural network approach than by the linear regression method.

The neural network Jacobians can provide not only an estimate of the mean sensitivity between two variables, but also an estimate of the distribution of the sensitivity. Figure 4 shows the histograms of estimated sensitivities for each pair of variables. Neural network sensitivity estimates agree well with the theoretical values as shown in the Jacobian matrix above. For example, from the theoretical Jacobian matrix, we know that the sensitivity of  $y_1$  to  $x_1$  is  $2x_1$ . This sensitivity has a normal distribution with a mean of 8 and variance equal to 4 as indicated in Fig. 4a. Similarly, the sensitivities of  $y_2$  to  $x_2$  are an exponential distribution with a mean of 1.7 (Fig. 4e). It is close to the theoretical distribution and mean, which is the same as those for  $y_2$ . As expected, the distributions of  $\partial y_3 / \partial x_1$  and  $\partial y_3 / \partial x_2$



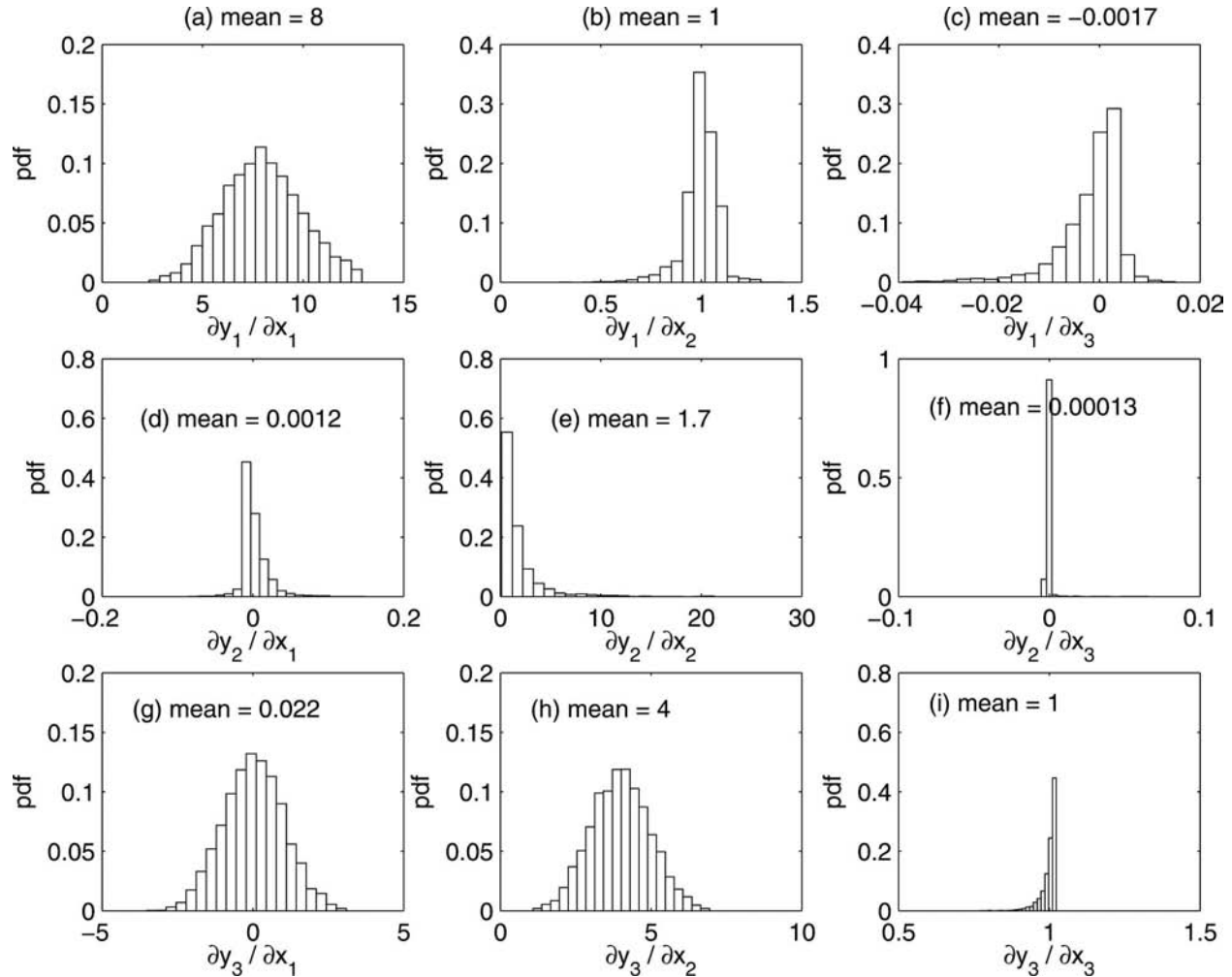


FIG. 4. Histograms of estimated sensitivities from the neural network for three test relationships. The mean sensitivity for each pair of variables is also given.

should be normal and given by  $N(0, 1)$  and  $N(4, 1)$ . The bottom panel in Fig. 4 indicates that the neural network sensitivities are close to the normal distributions for  $x_1$  and  $x_2$ .

#### 4. Relationships between CFL and cloud parameters

In this section, the relationships between CFL and other variables are discussed based on the characteristics of Jacobians from a new neural network approach. We use bivariate relationships only (in contrast with the previous section) to simplify the physical interpretation of the results. An NN is created for each of the cloud variables (the only input of the NN) and the CFL is the unique NN output. These four NNs are used to analyze the relationships for each pair of variables. We first plot

the neural network behavior curves against the scatter-plots for each pair; then we examine the histograms of the NN-estimated sensitivities.

##### a. Neural network behavior for pairs of variables

Figure 5 shows the root-mean-square errors (RMSEs) during the training process for each pair of variables. The RMSE decreases quickly to nearly constant by the end of the training. For CBH, the early stopping technique (Bishop 1996) is used to regularize the learning process because the RMSE from cross-validation data starts increasing. Given that the standard deviation of CFL is about  $26 \text{ W m}^{-2}$ , the LWP (RMSE  $\sim 15$ ) seems to best explain CFL variability, followed by CBT (RMSE  $\sim 19$ ), CLD (RMSE  $\sim 19$ ), and CBH (RMSE  $\sim 23$ ) in decreasing order.

The neural network behavior curves are examined to

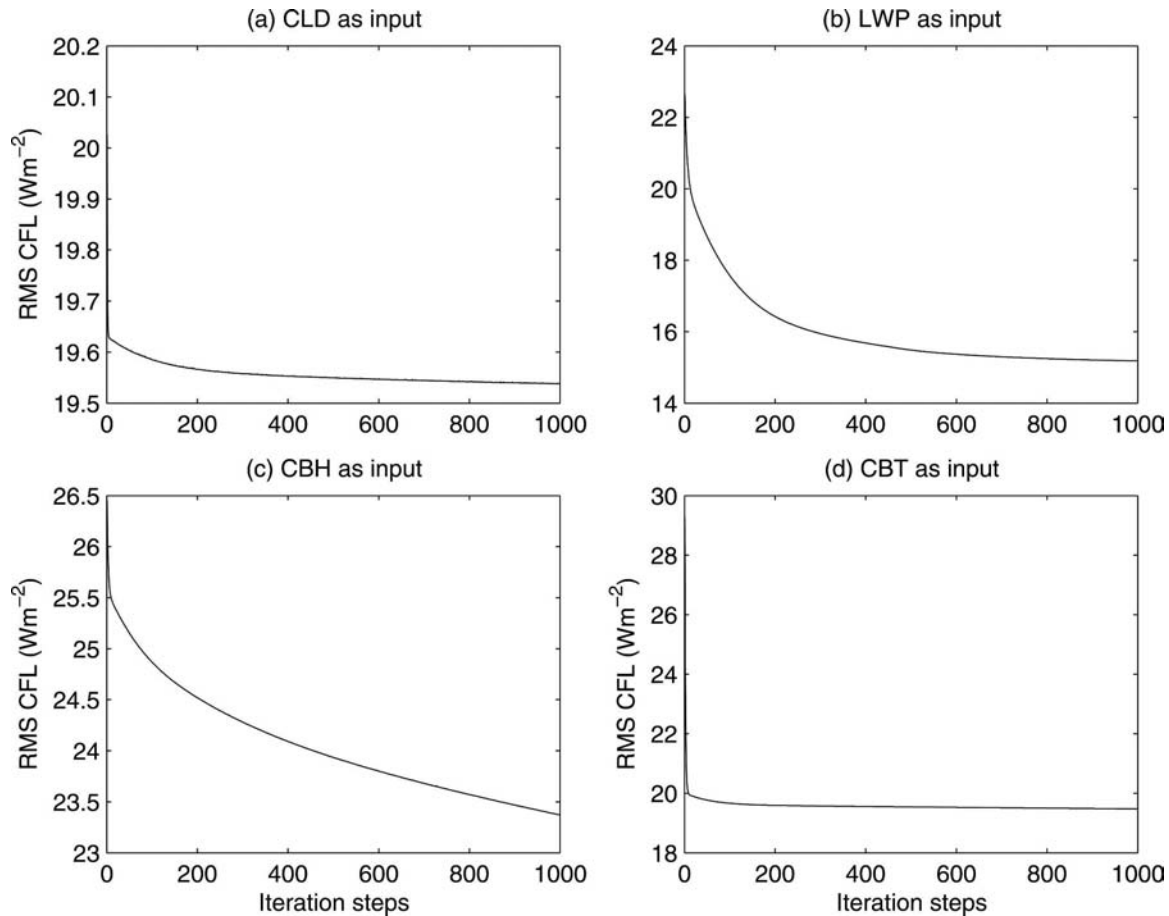


FIG. 5. The RMS error of CFL ( $\text{W m}^{-2}$ ) during the training process with a single input variable as (a) CLD, (b) LWP, (c) CBH (early stopping), and (d) CBT.

provide confidence in the sensitivity estimates in the following subsections. These curves are direct output from the trained neural network. The black lines in Fig. 6 are constructed by applying a range of input values to the trained neural network. The neural network behavior curve represents the averaged CFL corresponding to values of the cloud parameter. The limitation of the bivariate case is that much of the scatter cannot be explained by a single input variable; that is, interactions among the cloud variables contribute to the scatter. The slope of the curve at a given point is the sensitivity of CFL to the corresponding cloud parameter, and the steeper the curve, the larger the sensitivity. For example, when LWP values are low in Fig. 6b, the steep slope indicates a large sensitivity of CFL to LWP.

From the slopes of the neural network behavior curves in Fig. 6, we obtain a basic idea of the sensitivity characteristics. The sensitivity of CFL to cloud cover increases with cloud cover because the steepness of the behavior curve increases (Fig. 6a). For the sensitivity to

LWP, the largest slope occurs in the low range of LWP values, while the slope is small for high LWP values. This indicates a large sensitivity to low LWP values (Fig. 6b). Both negative and positive relationships exist between CFL and CBH. The slope is positive for low-altitude clouds, and negative for high-altitude clouds with a decrease in magnitude as CBH increases (Fig. 6c). The largest sensitivity between CFL and CBT occurs when base temperatures are neither extremely cold nor warm (Fig. 6d). The details of these sensitivities are discussed in following subsections.

Figure 7 shows the histograms of the Jacobians from an NN with CFL as output and each of the cloud variables (CLD, LWP, CBH, and CBT) as input. A bimodal distribution characterizes all of the Jacobian histograms for this dataset, albeit it is weak for  $\partial\text{CFL}/\partial\text{CBT}$ . This consistent pattern illustrates the nonlinear relationships between CFL and cloud properties, as well as the two distinct regimes that prevailed during SHEBA.



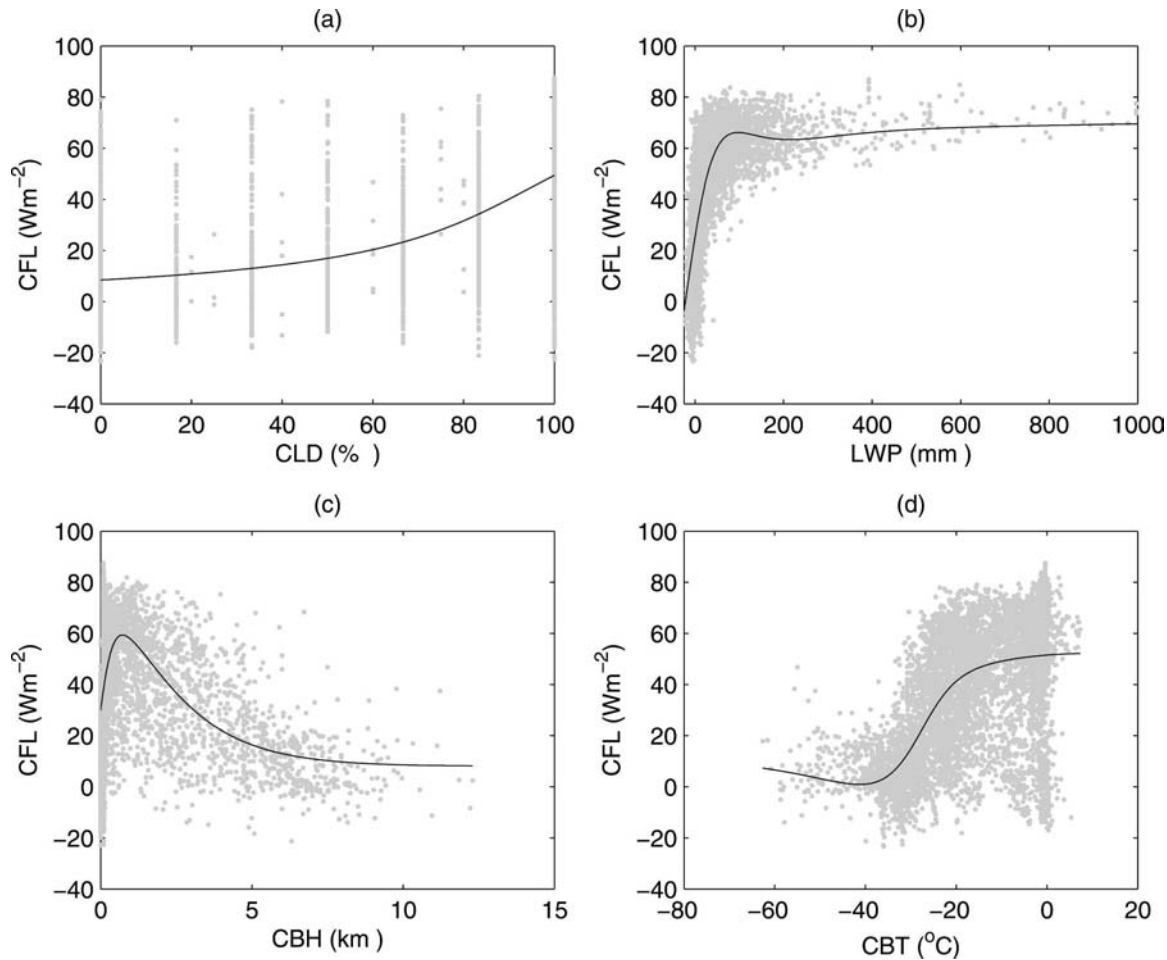


FIG. 6. CFL ( $\text{W m}^{-2}$ ) simulated with the NN using a single input variable: (a) CLD, (b) LWP, (c) CBH, and (d) CBT. The gray dots are data and the black curve is the NN estimate after training. The slope at each point along the curves represents the sensitivity of CFL to the corresponding cloud parameter. Steep slopes correspond to high sensitivities.

### b. CFL vs CLD

It is well known that the presence of cloud has a significant impact on the surface radiation budget in the Arctic. This is partly owing to the absence of solar energy in winter and partly to the prevalence of low-level temperature inversions. The lack of humidity in the Arctic atmosphere enhances the cloud impact. In the Tropics clouds have a weaker effect because of abundant low-level moisture. Using hourly measurements from SHEBA, the NN produces a sensitivity  $\partial\text{CFL}/\partial\text{CLD}$  ranging from 0.1 to 0.9 watts per square meter per percent cloudiness ( $\text{W m}^{-2} \%^{-1}$ ), with a mean of about  $0.68 \text{ W m}^{-2} \%^{-1}$ . Shupe and Intrieri (2004) show that  $\partial\text{CFL}/\partial\text{CLD}$  has a mean value of  $\sim 0.65 \text{ W m}^{-2} \%^{-1}$  with a range of 0.3 to  $0.8 \text{ W m}^{-2} \%^{-1}$ . These two independent results agree well. However, the bimodal distribution of the Jacobians indicates that the mean

value is not particularly meaningful (Fig. 7a) because this mean state does not exist often. The time scales are different in these two studies. Shupe and Intrieri (2004) obtain the sensitivity of CFL to CLD by looking at 2-day averages, while in the present study, hourly data are used. The same analysis using daily data produces similar results but with a slightly higher mean sensitivity. The NN approach, however, performs better and more reliably with a large dataset. Thus, we have more confidence in the realism of the results using hourly data.

To better understand the reason for the bimodal distribution in Fig. 7a, we examine separately the data points in the low and high sensitivity bins in this figure. For each of the two bins, we examine the distribution of CLD values and the predominant seasonal contributions to each bin. The two peaks in the sensitivity histogram in Fig. 7a correspond to the two regimes of

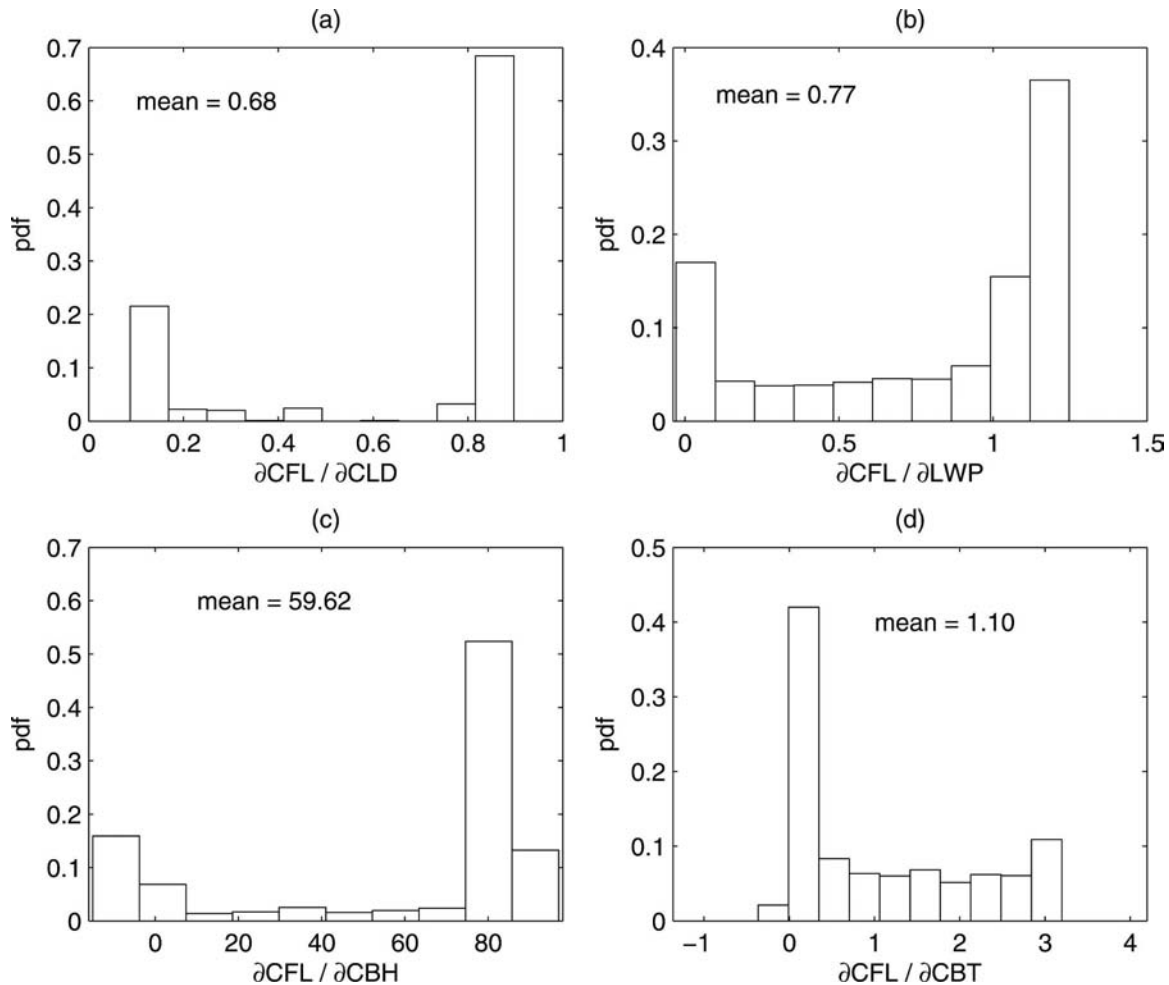


FIG. 7. Sensitivities between pairs of variables as represented by Jacobians from neural network for (a)  $\partial\text{CFL}/\partial\text{CLD}$  ( $\text{W m}^{-2} \%^{-1}$ ), (b)  $\partial\text{CFL}/\partial\text{LWP}$  ( $\text{W g}^{-1}$ ), (c)  $\partial\text{CFL}/\partial\text{CBH}$  ( $\text{W m}^{-2} \text{ km}^{-1}$ ), and (d)  $\partial\text{CFL}/\partial\text{CBT}$  ( $\text{W m}^{-2} \text{ C}^{-1}$ ). The number in each histogram is the mean sensitivity.

cloud presented in Fig. 8. The top (bottom) plots are for the low (high) sensitivity peak. The peak with low sensitivity corresponds to conditions with little or no cloud. These conditions occur mainly in winter (Figs. 8a and 8b). The high sensitivity peak contains overcast cases, which occur mostly in warm seasons (Figs. 8c and 8d). Shupe and Intrieri (2004) also point out that sensitivity increases with CLD.

The two distinct regimes with either clear or cloudy conditions have two implications. First, they indicate an asymmetric response to changes in cloud cover. The changes in CFL are plotted in Fig. 9 for transition cases: plus signs are for clear to cloudy, and circles for cloudy to clear. There is a slight tendency for the response to cooling (cloudy  $\rightarrow$  clear) to be stronger than that to warming (clear  $\rightarrow$  cloudy). The sample size, however, is too small to make a robust conclusion. The second implication is that the higher sensitivity during overcast

conditions may represent effects of other variables on CFL. For example, if there is more water vapor in the atmosphere ahead of an incoming cloud mass with a storm, it may have a significant effect on CFL. This combination of effects could be better represented by a multivariate NN model that uses interaction terms (this will be the subject of a future study).

### c. CFL vs LWP

Column LWP is used to represent cloud bulk microphysics. The NN produces a mean sensitivity  $\partial\text{CFL}/\partial\text{LWP}$  of  $0.78 \text{ W m}^{-2}$  per unit change in LWP (grams per square meter; hereafter  $\text{W g}^{-1}$ ). The bimodal distribution evident in Fig. 7b has a low-sensitivity peak near zero, and another high-sensitivity one near  $1.2 \text{ W g}^{-1}$ . Another group with a slightly lower sensitivity ( $1.1 \text{ W g}^{-1}$ ) has a comparable number of data points to those in the low-sensitivity peak group (hereafter re-

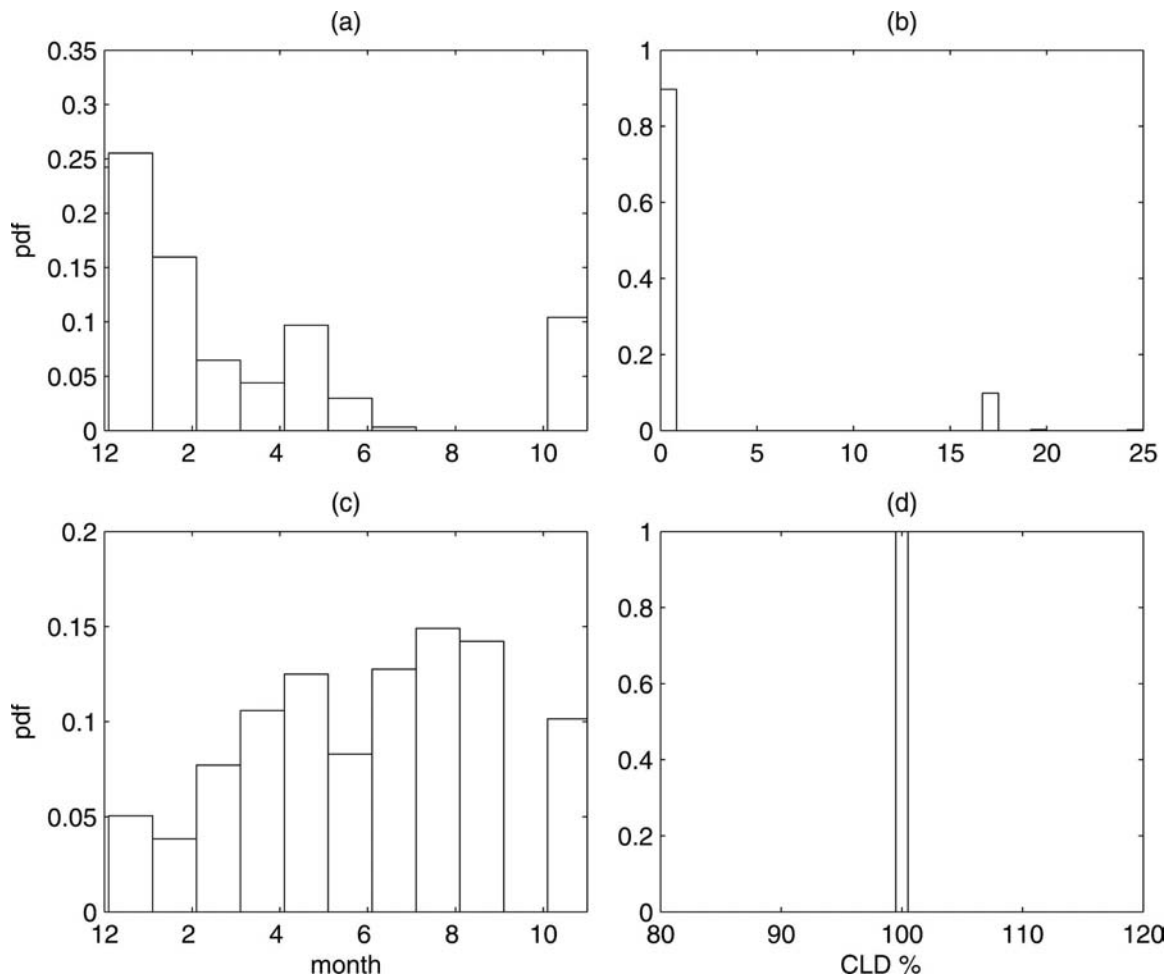


FIG. 8. Histograms showing monthly frequencies and corresponding distributions of cloud cover for low/high sensitivity  $\partial\text{CFL}/\partial\text{CLD}$  shown in Fig. 7a: (a), (b) for the low-sensitivity peak and (c), (d) for the high-sensitivity peak.

ferred to as bin 9). We focus on these three bins. LWP is large with a minimum value  $>80 \text{ g m}^{-2}$  in the low-sensitivity bin, which occurs mainly in late spring and summer (Fig. 10a). Thus, the impact of changes in LWP on CFL is near zero when LWP is high. This is known as the longwave saturation effect, which occurs because thick clouds with a large LWP emit as a blackbody, and any further change in LWP does not change CFL appreciably. These saturated cases occur mainly during summer. The high-sensitivity cases, on the other hand, occur when LWP is low (less than  $10 \text{ g m}^{-2}$ ), which occurs mainly during winter and spring (Fig. 10b). Clouds in bin 9 also have LWP values less than  $15 \text{ g m}^{-2}$ . All data in those two high-sensitivity bins are, however, within the range of retrieval uncertainty for LWP ( $\pm 25 \text{ W m}^{-2}$ ). Thus, we do not have enough confidence to define the LWP range for the high-sensitivity regimes.

The other bins in the Jacobian histogram (Fig. 7b)

indicate that the sensitivity  $\partial\text{CFL}/\partial\text{LWP}$  increases from 0.2 to  $1 \text{ W g}^{-1}$  as LWP decreases from 72 to  $6 \text{ g m}^{-2}$  (Fig. 11a). This negative relationship is consistent with the findings of Shupe and Intrieri (2004), and illustrates the longwave saturation effect. The empirical relation between cloud emissivity and LWP is plotted in Fig. 11b based on Stephens's (1978) parameterization with a total mass absorption coefficient equal to  $0.158 \text{ m}^2 \text{ g}^{-1}$ . As LWP increases, cloud emissivity quickly approaches that of a blackbody. Thus, further increases in LWP do not affect the downward longwave flux.

#### d. CFL vs CBH

CBH does not have a distinct seasonal trend (Intrieri et al. 2002a), nor does it have a direct impact on CFL. However, it does have an indirect impact because the CBH affects the temperature of the cloud base, which has a direct effect on the surface radiation budget. This is particularly important because low-level Arctic tem-

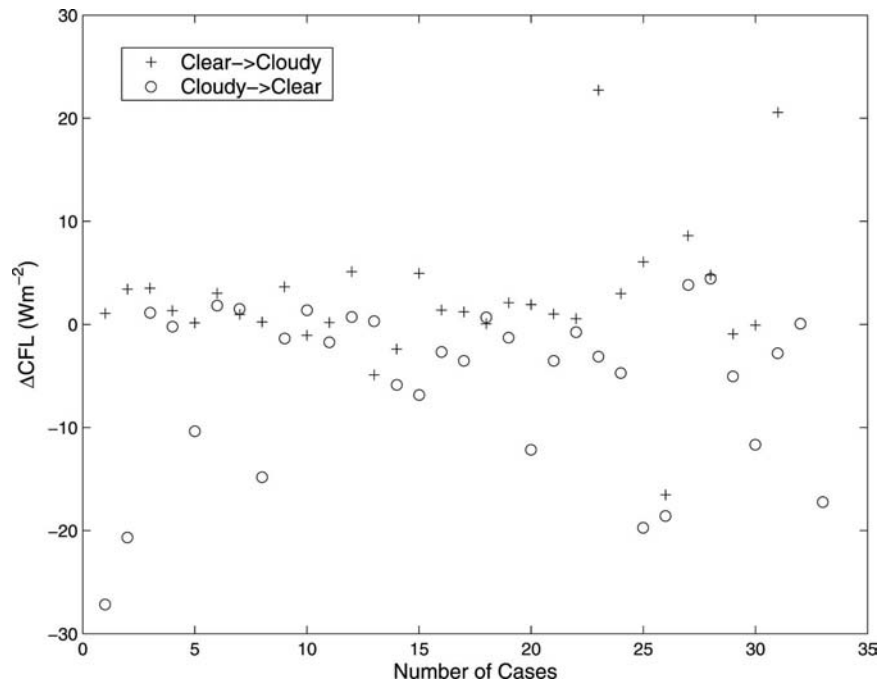


FIG. 9. Changes in CFL for transitions from clear to overcast (crosses) and from overcast to clear (circles) conditions.

perature inversions are frequent, especially in winter. There are both positive and negative relationships between CFL and CBH, as indicated by positive and negative sensitivities ( $\partial\text{CFL}/\partial\text{CBH}$ ) in Fig. 7c. This likely indicates the different effects of clouds residing in or above the temperature inversion layer.

The peak with negative sensitivities corresponds to a group of clouds with a minimum CBH of 0.8 km (Figs. 12a,b). Most of these clouds are above 1 km, which is

about the average height of the temperature inversion in the Arctic. Above this height, CBT generally decreases as CBH increases; thus, CFL decreases. The peak with high positive sensitivities corresponds to a group of very low clouds (less than 0.16 km) (Figs. 12c,d). The bin with the highest positive sensitivity is characterized by zero CBH. This indicates that changes in clouds extending to the ground (or fog), especially if they contain liquid water, have a large impact on CFL.

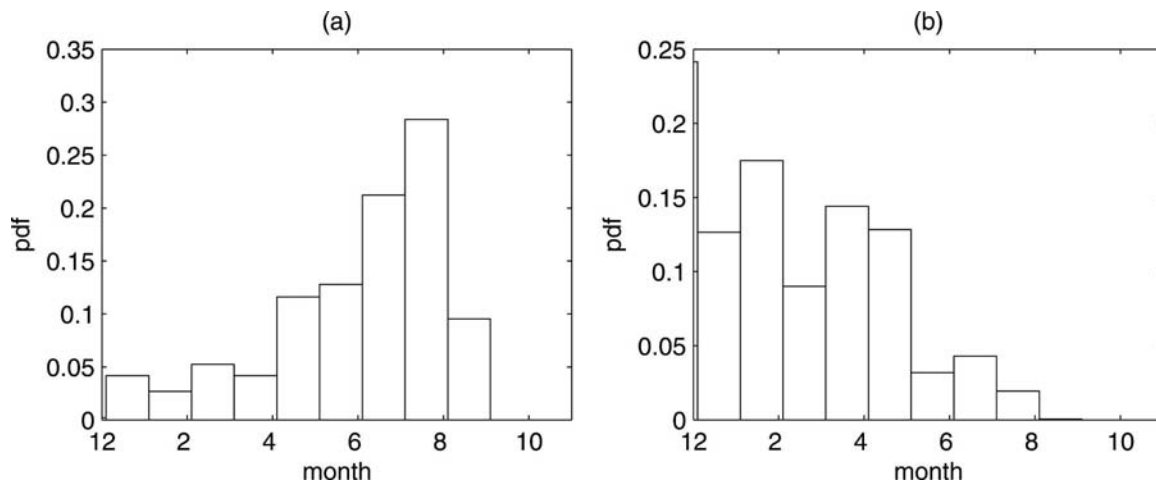


FIG. 10. Histograms showing monthly frequencies of LWP for low/high sensitivity  $\partial\text{CFL}/\partial\text{LWP}$  shown in Fig. 7b: (a) for the low-sensitivity peak and (b) for the high-sensitivity peak.

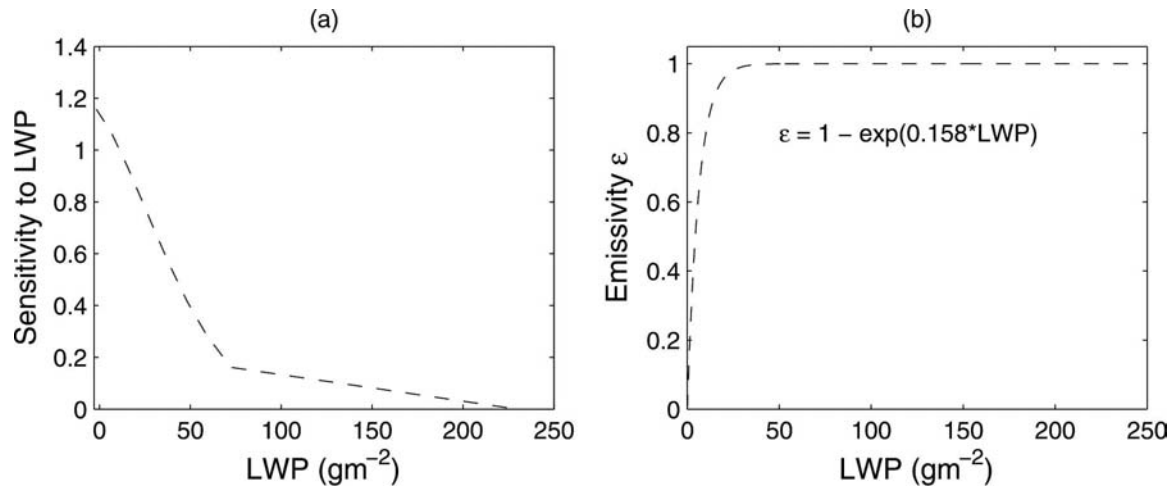


FIG. 11. The longwave saturation effect as indicated in (a) observed sensitivity of CFL to LWP as a function of LWP; (b) empirical relationship between cloud emissivity  $\epsilon$  and LWP based on Stephens (1978).

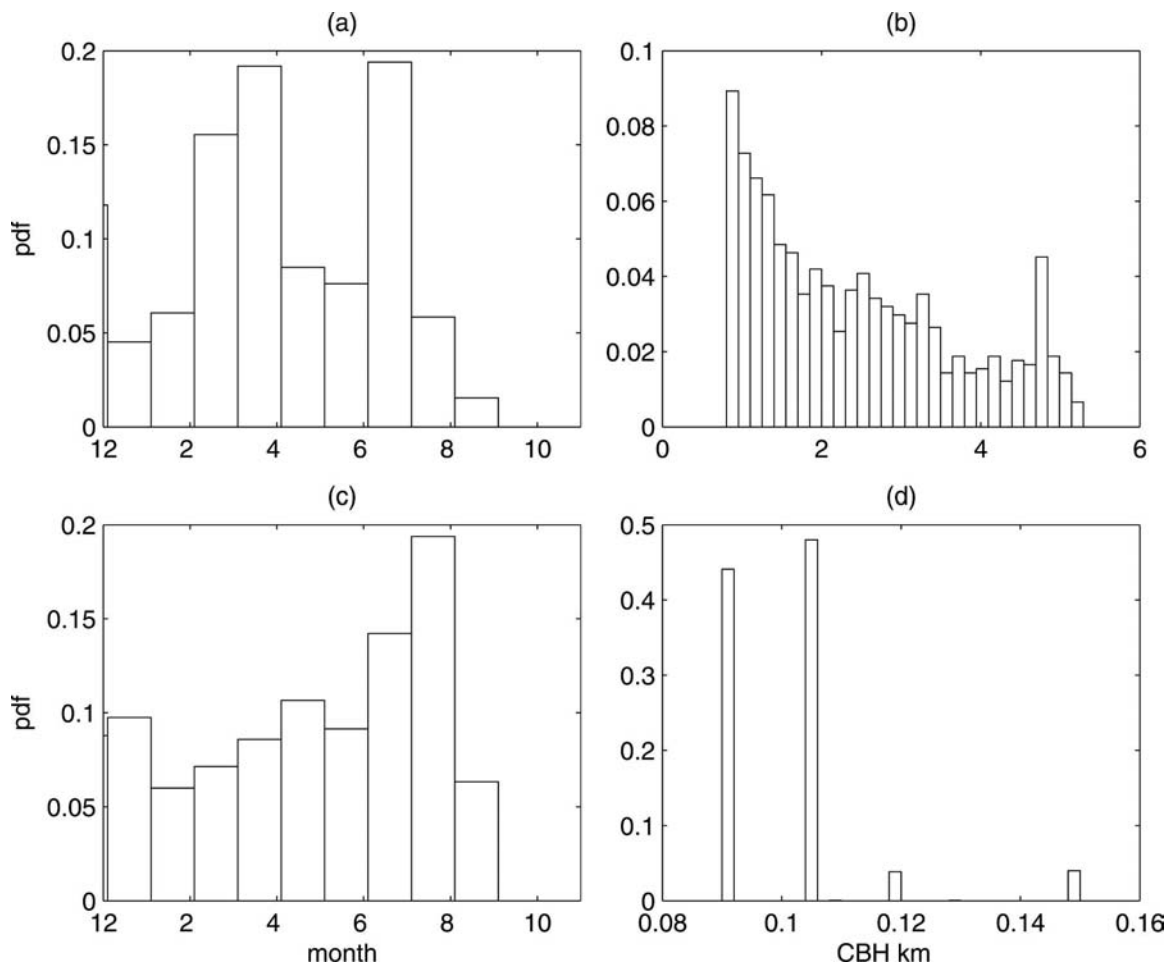


FIG. 12. Histograms showing monthly frequencies and corresponding distributions of CBH for low–high sensitivity  $\partial\text{CFL}/\partial\text{CBH}$  shown in Fig. 7c: (a), (b) for the low-sensitivity peak and (c), (d) for the high-sensitivity peak.

TABLE 1. Sensitivities  $s$  of CFL to CBH from the NN.

	$\partial\text{CFL}/\partial\text{CLD} \text{ (W m}^{-2} \text{ km}^{-1}\text{)}$			
	$s < 10$	$10 < s < 0$	$0 < s < 10$	$s > 10$
CBH (km)	1.02–3.17	$>3.20$ or $0.75 < \text{CBH} < 1.00$	0.60–0.72	$<0.57$

These clouds are also the most difficult to detect and measure by satellite sensors. This research emphasizes the importance of improving retrieval algorithms for low Arctic clouds.

Shupe and Intrieri (2004) divided clouds into three groups to examine the relationships between CFL and CBH: clouds with  $\text{CBH} < 0.5 \text{ km}$ , clouds with  $\text{CBH} > 3 \text{ km}$ , and clouds in between. The neural network produces a similar but more detailed division of clouds (Table 1). The resolution of CBH measurements is  $0.029 \text{ km}$ . The clouds with large negative sensitivities

have bases between 1 and 3 km. Clouds with large positive sensitivities have a base height lower than about  $0.60 \text{ km}$ . When clouds are near but below the inversion layer, or above  $3 \text{ km}$ , the sensitivity is smaller.

*e. CFL vs CBT*

CBT has a significant effect on the radiative properties of clouds. It directly affects downward longwave flux to the surface, and thus CFL at the surface. Warm clouds (high CBT) are responsible for the largest CFL (Shupe and Intrieri 2004). However, the NN Jacobian

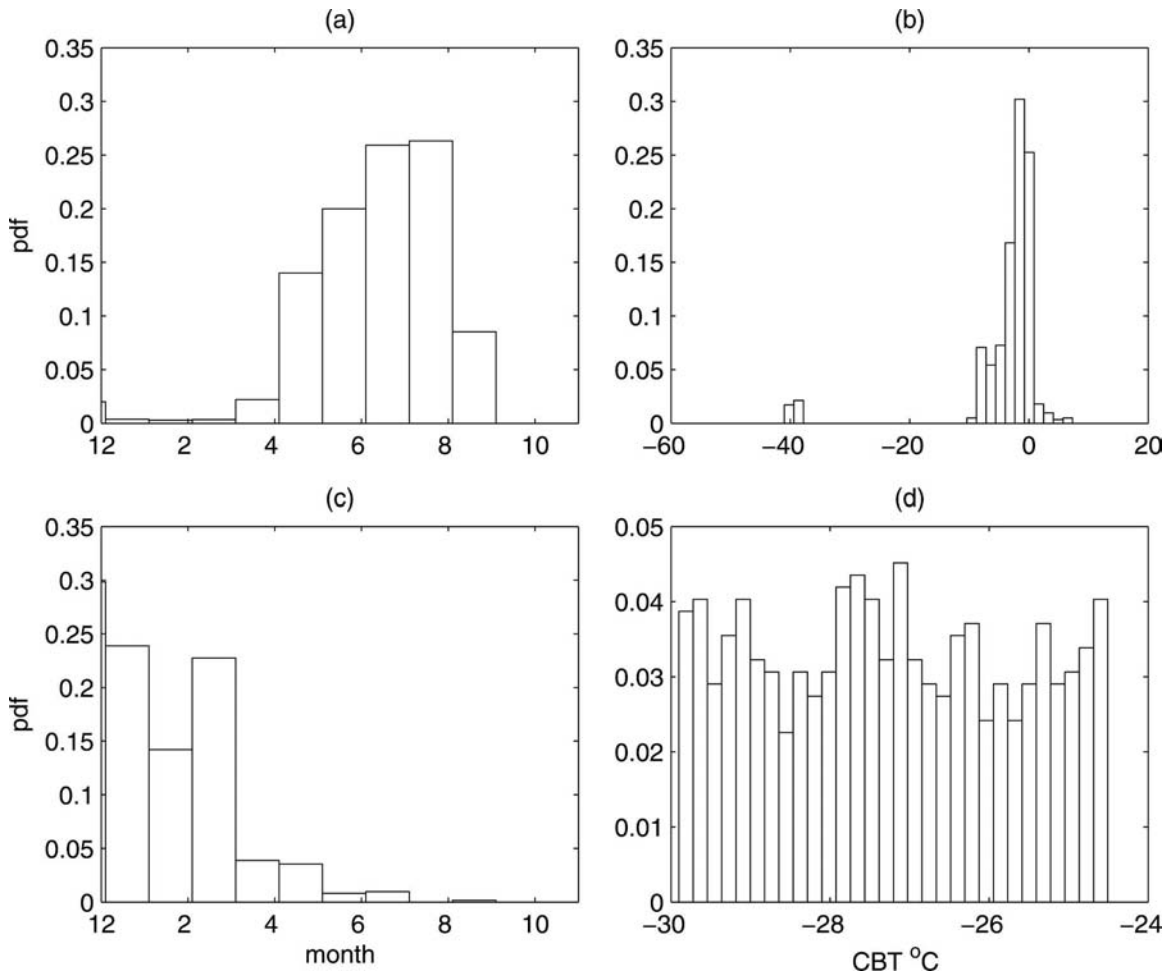


FIG. 13. Histograms showing monthly frequencies and corresponding distributions of CBT for low–high sensitivity  $\partial\text{CFL}/\partial\text{CBT}$  shown in Fig. 7d: (a), (b) for the low-sensitivity peak and (c), (d) for the high-sensitivity peak.



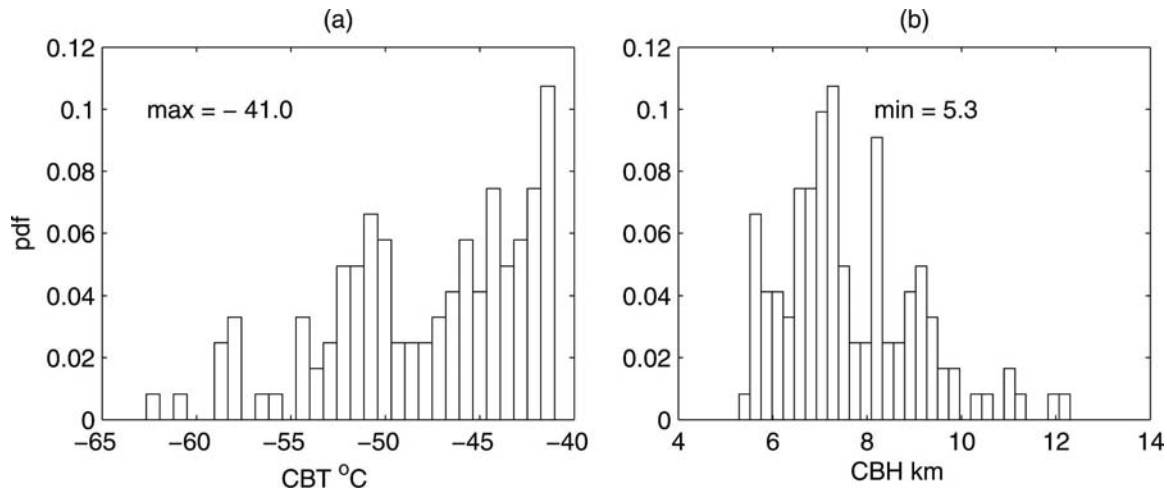


FIG. 14. Data distribution in terms of (a) CBT and (b) CBH corresponding to the bin with near-zero sensitivities in the histogram for  $\partial\text{CFL}/\partial\text{CBT}$  in Fig. 7d.

shows that the sensitivities for these warm clouds are not necessarily large, as shown by the peak with low sensitivity in the Jacobian histogram (Fig. 7d). This peak corresponds mainly to a group of clouds with base temperatures higher than  $-9^{\circ}\text{C}$  (Fig. 13b), which appear during spring and summer (Fig. 13a). Another group of clouds with sensitivities near zero, as indicated by the first bin in the histogram (Fig. 7d), has very low base temperatures (less than  $-40^{\circ}\text{C}$ ). This group occurs infrequently, and is characterized by high clouds with a minimum height above 5 km (Fig. 14). The highest sensitivity of CFL to CBT occurs during winter and early spring (Fig. 13c) with an average range of CBT between  $-24^{\circ}$  and  $-30^{\circ}\text{C}$  (Fig. 13d). Most clouds in this group have a low base height and exist under the temperature inversion. This causes the CFL to be highly sensitive to changes in CBH and CBT. These results show that clouds within or near the inversion have the largest impact on CFL. The NN produces a mean sensitivity  $\partial\text{CFL}/\partial\text{CBT}$  of  $1.10 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$  (Fig. 7d), which is consistent with the average value of about  $1 \text{ W m}^{-2} \text{ }^{\circ}\text{C}^{-1}$  found by Shupe and Intrieri (2004) under typical Arctic conditions.

## 5. Summary and discussion

The neural network is able to capture the sensitivity of CFL to a wide variety of cloud properties including cloud cover, LWP, CBH, and CBT. The bimodal distribution in sensitivity of CFL with respect to each of the other four variables clearly shows the nonlinear behavior in the relationships, and provides a wealth of additional information for interpreting them. This novel approach is able to characterize different “re-

gimes” in the system without a priori assumptions. It should be noted that the instantaneous longwave surface radiative cloud forcing is also influenced by the transmissivity of the atmosphere below the cloud; thus, the CLF of a high cloud will be more susceptible to this effect than will a low cloud. The Arctic atmosphere typically contains little moisture, however, so the transmissivity of the below-cloud layer will be of less importance than at lower latitudes. We also recognize that over time the emission from the cloud will likely warm the atmosphere below it, thereby adding to the CLF.

Table 2 gives a summary of the Jacobians from a neural network using pairs of variables at an hourly time scale. The overall mean sensitivities are listed as are the averages for each mode and corresponding cloud characteristics. The strongest bimodal case is for  $\partial\text{CFL}/\partial\text{CLD}$ . This bimodal sensitivity pattern matches the bimodality of cloud cover (Fig. 1b); that is, the response of the CFL is different under clear and cloudy conditions. Although the definition of cloud cover here is somewhat different from traditional ones (see section 2), the bimodality of Arctic cloud has been shown in a previous study using Russian drifting station data (Walsh and Chapman 1998). The mean sensitivity state of  $0.65 \text{ W m}^{-2} \%^{-1}$  agrees well with Shupe and Intrieri (2004), but our study indicates that it rarely exists in reality.

CFL increases as liquid water increases, but the slope of the relationship decreases for large LWP. Ultimately, it reaches a saturation state, in which additional increases in LWP have little effect on CFL. For example, the low-sensitivity regime is characterized by clouds with LWP larger than  $80 \text{ g m}^{-2}$  (Table 2). Those clouds often occur during summer with high CBT. The

TABLE 2. Sensitivities between pairs of variables represented as Jacobians from the neural network. Modes 1 and 2 correspond to low- and high-sensitivity peaks, respectively, in Fig. 7. For  $\partial\text{CFL}/\partial\text{CBH}$ , modes 1 and 2 correspond to negative- and positive-sensitivity peaks, respectively.

Jacobians		Mode 1 (low sensitivity)	Mean sensitivity	Mode 2 (high sensitivity)
$\partial\text{CFL}/\partial\text{CLD}$	$\text{W m}^{-2} \%^{-1}$	0.10	0.65*	0.90
	Conditions	Clear sky ( $\text{CLD} < 5\%$ )		Cloudy sky ( $\text{CLD} > 95\%$ )
	Primary season	Winter		Spring and summer
$\partial\text{CFL}/\partial\text{LWP}$	$\text{W g}^{-1}$	0.01	0.78	1.18
	Conditions	$\text{LWP} > 80 \text{ g m}^{-2}$		$\text{LWP} < 10 \text{ g m}^{-2}$
	Primary season	Summer		Winter
$\partial\text{CFL}/\partial\text{CBH}$	$\text{W m}^{-2} \text{ km}^{-1}$	−10.07	59.62	83.25
	Conditions	$\text{CBH} > 0.8 \text{ km}$		$\text{CBH} < 0.16 \text{ km}$
	Primary season	—		—
$\partial\text{CFL}/\partial\text{CBT}$	$\text{W m}^{-2} ^\circ\text{C}^{-1}$	0.19	1.10*	3.07
	Conditions	$\text{CBT} > -9^\circ\text{C}$		$\text{CBT} -24^\circ \text{ to } -30^\circ\text{C}$
	Primary season	Summer		Winter and early spring

\* Consistent with the results of Shupe and Intrieri (2004).

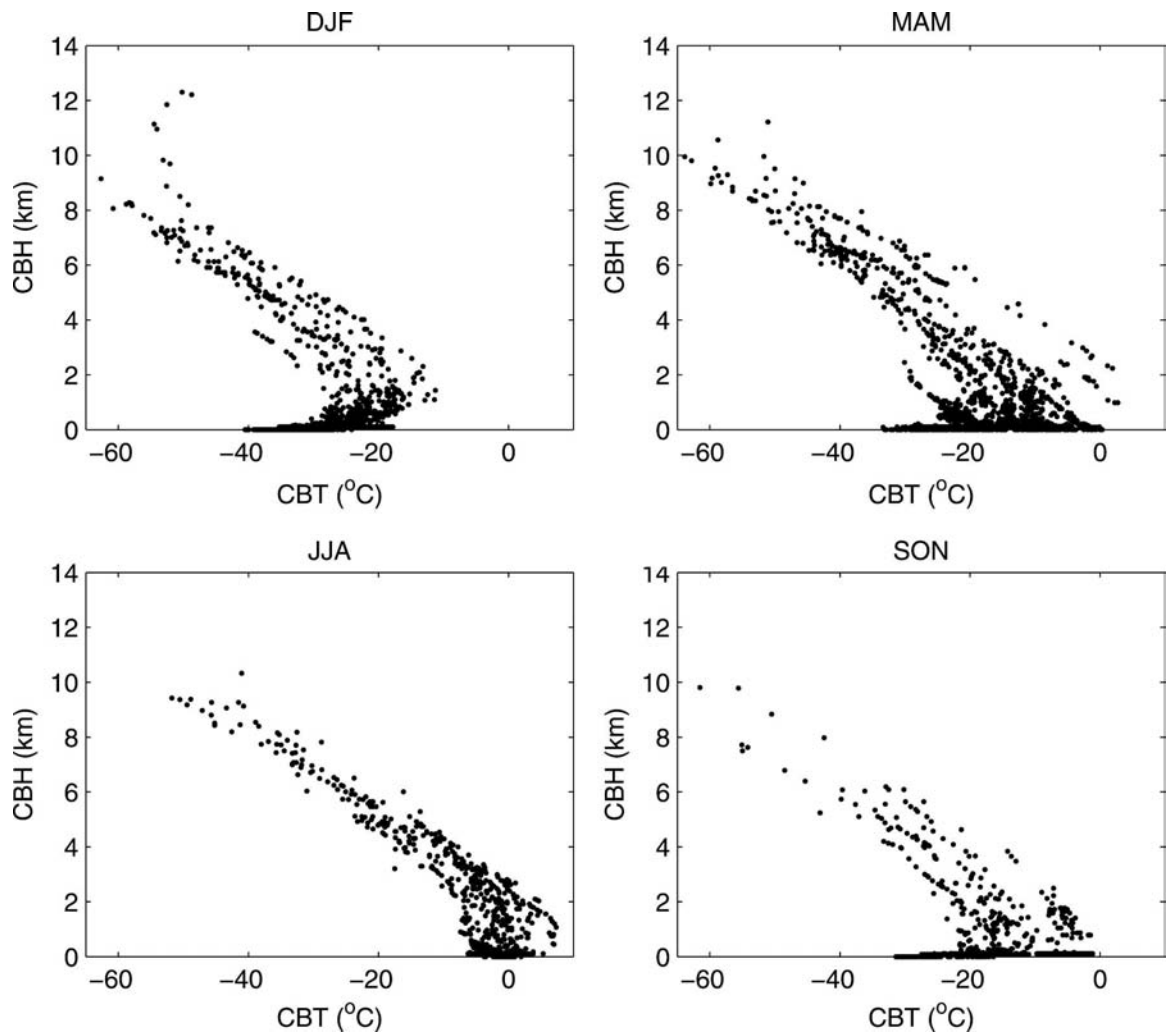


FIG. 15. Scatterplots of CBH vs CBT for winter [Dec–Jan–Feb (DJF)], spring [Mar–Apr–May (MAM)], summer [Jun–Jul–Aug (JJA)], and autumn [Sep–Oct–Nov (SON)].

CFL is most sensitive to clouds with very small LWP, which are within the uncertainty of the LWP retrievals. This suggests that further study of the relationship between CFL and LWP requires more accurate LWP measurements.

CBH and CBT are closely related; thus, their impacts on CFL are interrelated. Evidence of the near-surface temperature inversion is indicated in the scatterplot between CBH and CBT (Fig. 15), especially in winter. A strong negative relationship exists for high-altitude clouds, and a slightly positive one for low-altitude clouds. The negative correlation is dominant during summer when temperature inversions are much weaker.

When clouds reside within the temperature inversion layer, the relationship between CFL and CBH is positive. When CBH is higher than 0.8 km, CFL has a general tendency to decrease as CBH increases (Table 2). A small group of very cold, high ( $>5$  km) clouds corresponds to very low sensitivity of  $\partial\text{CFL}/\partial\text{CBT}$  (Fig. 14). This is a saturation effect similar to that found for CFL and LWP; that is, the sensitivity of CFL to CBT is small for very warm clouds ( $\text{CBT} > -9^{\circ}\text{C}$ ). This also occurs mainly in summer.

The sensitivities of CFL to cloud properties have been discussed by Shupe and Intrieri (2004) using SHEBA data. The mean states for sensitivities obtained in their study are in good agreement with ours derived with the NN model. The bimodal distribution, however, indicates that the mean cloud conditions rarely exist, and thus have little practical meaning. Shupe and Intrieri (2004) also discussed the different sensitivity values corresponding to cloud groups with different properties. Their results are based on the foreknowledge of how to separate different cloud groups, while the NN approach requires no a priori information.

Our study shows that the neural network can readily separate the different response regimes. This is encouraging because all results are derived from data without other subjective knowledge. Although the bivariate case cannot explain the total variability in CFL owing to the interaction of cloud variables, it demonstrates an effective way to quantify the relationships between climate variables. The next step will be to perform a multivariate study to understand the effects of cloud-variable interactions on sensitivities and to obtain multiple sensitivities simultaneously. This neural network/Jacobian method can also be applied to other groups of variables from measurements, retrievals, or model simulations to elucidate complex, nonlinear relationships that are prevalent in the climate system.

**Acknowledgments.** This work was supported by Grant NAG5-11720 from the National Aeronautics and

Space Administration and by funding from the Institute of Marine and Coastal Sciences, Rutgers University. We are thankful to Drs. John A. Beesley, Jean-Louis Dufresnes, and John Wilkin, and the anonymous reviewers for their helpful comments. We also thank Matthew Shupe for providing SHEBA data.

## REFERENCES

- Aires, F., and W. B. Rossow, 2003: Inferring instantaneous, multivariate and nonlinear sensitivities for the analysis of feedback processes in a dynamical system: The Lorenz model case study. *Quart. J. Roy. Meteor. Soc.*, **129**, 239–275.
- , M. Schmitt, N. Scott, and A. Chedin, 1999: The weight smoothing regularisation for MLP for resolving the input contribution's errors in functional interpolations. *IEEE Trans. Neural Networks*, **10**, 1502–1510.
- , C. Prigent, W. B. Rossow, and M. Rothstein, 2001: A new neural network approach including first-guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations. *J. Geophys. Res.*, **106**, 14 887–14 907.
- , —, and —, 2004: Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 3. Network Jacobians. *J. Geophys. Res.*, **109**, D10305, doi:10.1029/2003JD004175.
- Bishop, C. M., 1996: *Neural Networks for Pattern Recognition*. Oxford University Press, 482 pp.
- Chevallier, F., and J. F. Mahfouf, 2001: Evaluation of the Jacobians of infrared radiation models for variational data assimilation. *J. Appl. Meteor.*, **40**, 1445–1461.
- Comiso, J. C., 2003: Warming trends in the Arctic from clear sky satellite observations. *J. Climate*, **16**, 3498–3510.
- Davis, D., Z. Chen, L. Tsang, J. Hwang, and A. Chang, 1993: Retrieval of snow parameter by iterative inversion of a neural network. *IEEE Trans. Geosci. Remote Sens.*, **31**, 842–852.
- Escobar, J., A. Chédin, F. Chérut, and N. A. Scott, 1993: Réseaux de neurones multicouches pour la restitution de variables thermodynamiques atmosphériques à l'aide de sondeurs verticaux satellitaires. *Rend. Acad. Sci.*, **317B**, 911–918.
- Gardner, M. W., and S. R. Dorling, 2001: Artificial neural network-derived trends in daily maximum surface ozone concentrations. *J. Air Waste Manage.*, **51**, 1202–1210.
- Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feed-forward networks are universal approximators. *Neural Networks*, **2**, 359–366.
- Intrieri, J. M., M. D. Shupe, T. Uttal, and B. J. McCarty, 2002a: An annual cycle of Arctic cloud characteristics observed by radar and lidar at SHEBA. *J. Geophys. Res.*, **107**, 8030, doi:10.1029/2000JC000423.
- , C. F. Fairall, M. D. Shupe, O. G. P. Persson, E. L. Andreas, P. Guest, and R. M. Moritz, 2002b: An annual cycle of Arctic surface cloud forcing at SHEBA. *J. Geophys. Res.*, **107**, 8039, doi:10.1029/2000JC000439.
- Key, J. R., and A. J. Schweiger, 1998: Tools for atmospheric radiative transfer: Streamer and FluxNet. *Comput. Geosci.*, **24**, 443–451.
- Krasnopolsky, V. M., and F. Chevallier, 2003: Some neural network applications in environmental sciences. Part II: Advancing computational efficiency of environmental numerical models. *Neural Networks*, **16**, 335–348.
- Schweiger, A. J., 2004: Changes in seasonal cloud cover over the

- Arctic seas from satellite and surface observations. *Geophys. Res. Lett.*, **31**, L12207, doi:10.1029/2004GL020067.
- Serreze, M. C., and Coauthors, 2000: Observational evidence of recent change in the northern high-latitude environment. *Climatic Change*, **46**, 159–207.
- Shupe, M. D., and J. M. Intrieri, 2004: Cloud radiative forcing of the Arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle. *J. Climate*, **17**, 616–628.
- Stephens, G. L., 1978: Radiation profiles in extended water clouds. II: Parameterization schemes. *J. Atmos. Sci.*, **35**, 2123–2132.
- , 2005: Cloud feedbacks in the climate system: A critical review. *J. Climate*, **18**, 237–273.
- Thiria, S., C. Meijia, and F. Badran, 1993: A neural network approach for modeling nonlinear transfer function: Application for wind retrieval from spaceborne scatterometer data. *J. Geophys. Res.*, **98**, 22 827–22 841.
- Walsh, J. E., and W. L. Chapman, 1998: Arctic cloud–radiation–temperature associations in observational data and atmospheric reanalyses. *J. Climate*, **11**, 3030–3045.
- Wang, X., and J. R. Key, 2003: Recent trends in Arctic surface, cloud and radiation properties from space. *Science*, **299**, 1725–1728.
- Westwater, E. R., Y. Han, M. D. Shupe, and S. Y. Matrosov, 2001: Analysis of integrated cloud liquid and precipitable water vapor retrievals from microwave radiometers during SHEBA. *J. Geophys. Res.*, **106**, 32 019–32 030.
- Yi, J., and R. Prybutok, 1996: A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialised urban area. *Environ. Pollut.*, **92**, 349–357.